# Introduction

Enverus, an energy technology company, serves as the driving force behind this research endeavor. Since its establishment in 1999, Enverus has emerged as a prominent provider of energy market data, analytics, and technology solutions. With a commitment to optimizing operations and fostering a deeper understanding of energy markets, Enverus offers innovative software, data, and services to facilitate informed decision-making within the energy sector. Indeed, Enverus gives energy firms the platforms, tools, and applications they need to be adaptable and successful in a challenging and changing market environment. Additionally, Enverus provides invaluable services such as expert guidance, data analysis, and market intelligence, further solidifying its position as a leader in the field.

This project aims to investigate the methodologies utilized for predicting the performance of solar farms, with a specific focus on the geographical impact and modeling techniques employed. Our team will experiment on different models with distinct parameters in order to facilitate a comprehensive comparison and contrast of various modeling scenarios. This project particularly inspires to combine various data sources and develop specific modeling techniques to explore the implications for solar farm performance prediction.

The problem statement at the core of this project revolves around comparing and contrasting different design options concerning the utilization of location-specific source data and alternative modeling techniques. Enverus will contribute anonymized data sets encompassing input and target variables, while also suggesting a range of modeling techniques to be explored. The design matrix adopted for this study incorporates a variety of data sources, ranging from macro to meso and micro regions in proximity to the area under investigation, as well as multiple time-series models including Linear Regression, ARIMA, or XGBoost. The project's overarching objective is to be able to answer the following hypothesis questions.

Hypothesis questions:

* When comparing the use of generic data sources and traditional modeling methods to solar farm performance prediction, how does the integration of location-specific data sources and sophisticated modeling techniques affect accuracy and predictive performance?
* Moreover, how might these enhancements help energy firms maximize operations, allocate resources, and pinpoint areas with the greatest potential for solar power generation?

The findings are expected to gain valuable insights into the renewable energy industry and aid in the development of more accurate and efficient modeling techniques. Ultimately, this project endeavors to facilitate informed decision-making in the planning and implementation of solar farms, promoting the adoption of sustainable energy solutions and furthering the advancement of the renewable energy sector.

# Literature survey

Solar power generation is affected by several environmental factors. These include factors such as solar irradiance, temperature, humidity, dust, shading, and wind speed (Singh, & Singh). In addition, technical design features of the photovoltaic cells, such as the material used in manufacturing the PV cells, affect the cell power generation (Chikate et.al).

# Methodology

## 3.1 Data Collection

## 3.2 Preprocessing

## 3.3 Data Modelling

## 3.4 Experimentation Design

# Results and Visualizations

# Conclusion and discussion

References

A. K. Singh and R. R. Singh, "An Overview of Factors Influencing Solar Power Efficiency and Strategies for Enhancing," 2021 Innovations in Power and Advanced Computing Technologies (i-PACT), Kuala Lumpur, Malaysia, 2021, pp. 1-6, doi: 10.1109/i-PACT52855.2021.9696845.

Chikate, Bhalachandra V., Y. Sadawarte, and B. D. C. O. E. Sewagram. "The factors affecting the performance of solar cell." *International journal of computer applications* 1.1 (2015): 0975-8887.